# Inadmissible Minimax Invariant Estimator (統計的決定函数のAdmissibilityの研究報告集)



石井, 吾郎

数理解析研究所講究録
27巻 1967年 45-60



# Inadmissible minimax invariant estimator

石井吾郎（大阪市大．商）

## §1 序

$\mathscr{X} = \{x\}$ ：　標本空間

$\Theta = \{\theta\}$ ：　母数空間

$\mathscr{P} = \{ f(x,\theta),\ \theta \in \Theta \}$ ：$\mathscr{X}$上の確率分布族．$f$は密度関数

$A = \{a\}$ ：　行動空間

$\Delta = \{\delta\}$ ; $\delta : \mathscr{X} \to A$ への map，決定関数

$L(\theta, a)$ ：　損失関数

$\rho(\theta, \delta) = E_\theta \left( L(\theta, \delta(x)) \right)$ ：　リスク．

$G = \{g\}$ ：　$\mathscr{X}$上の変換群

$\mathring{G} = \{\mathring{g}\}$ ：　$\Theta$上の変換群

$\bar{G} = \{\bar{g}\}$ ：　$A$上の変換群

$G,\ \mathring{G},\ \bar{G}$ 相互の関係は $g \longleftrightarrow \mathring{g}$ 同型，$\mathring{g} \to \bar{g}$ 準同型とする

$L(\mathring{g}\theta, \bar{g}a) = L(\theta, a)$　のとき損失は$G$－不変であるという．

$\delta(gx) = \bar{g}\,\delta(x)$　のとき $\delta$ を$G$－不変な決定関数という．

$\Delta_G = \{ \delta : \delta(gx) = \bar{g}\,\delta(x) \}$ ：$G$－不変な決定関数全体

損失関数がある変換群$G$について不変であるときには$\Delta_G$の中からなるべくよい$\delta$を求めると云う原則を立てて見る。そのとき イ $\delta$ が$\Delta_G$の中でミニマックスであれば$\Delta$の中でミニマックスであろう　とか ロ $\delta$ が$\Delta_G$の中で許容的であれば$\Delta$の中で許容的であろう　という様な期待がありうるわけであります。実際そうなっている例は沢山あるし、又適当な

条件の下で①, ② が成立することを示す定理もあります。Kudo. Nat. Sci. Rep. Ochanomizu Uni. 1955. Kiefer. Ann. Math. Stat. 1957. 適当な条件の下で肯定的な結論の出る場合の話はこれらの論文あるいは次の Stein の論文の引用文献を見て頂くことにしまして、ここで紹介しようとしますことは

Stein [1] 3rd Berkeley Symposium

[2] 4th Berkeley Symposium

[3] Hotelling 記念論文集

にあります否定的な例であります。則ち $\Delta_G$ の中から良いものを選んで見たらそれより良いものが $\Delta$ の中で見つけられると云う例であります。

次節で Stein [1], [2] の結果を少し修正した形で述べます。

## §2 最小2乗推定は許容的でない ($s>2$ のとき)

観測値 X は次の線形構造をもっていると仮定する。

(1) $\quad X = A\mu + \varepsilon$

但し $\quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $\quad A = \begin{bmatrix} a_{11} & \cdots & a_{1s} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{ns} \end{bmatrix}$, $\quad \mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_s \end{bmatrix}$, $\quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$

$$n > s \geq 3,\ \mathrm{rank}\, A = s, \quad \varepsilon : N(0, I_n\sigma^2)$$

$A$ の列ベクトルで張られる $R^n$ の部分空間を $L(A)$ で示す。$R^n$ より $L(A)$ への射影作用素を $e$ とする。

$$e = A(A'A)^{-1}A', \quad \mathrm{rank}\, e = s$$

$\mu$ の最小2乗推定を $\hat{\mu}(X)$ とすると、それは $X$ より $L(A)$ への射影の足で与

えられ

(2) $\hat{\mu}(x) = (A'A)^{-1}A'X$

$$\hat{\mu} \quad : \quad N(\mu, \ \Sigma\sigma^2) \quad , \quad \Sigma = (A'A)^{-1}$$

である。

(3) $S(x) = X'(I_n - e)X$

とおくと，$S$ と $\hat{\mu}$ は独立で $S/\sigma^2$ は自由度 $n-s$ のカイ２乗分布に従う。

$$E(S) = (n-s)\sigma^2 \quad , \quad E(S^2) = (n-s)(n-s+2)\sigma^4$$

$\theta = (\mu, \sigma^2) \in \Theta = R^s \oplus R^+$ が母数及び母数空間で

$(\hat{\mu}, S)$ は最小十分統計量である。

$a \in R^+, \quad b \in R^s$ に対し

$$g_{ab} : X \longrightarrow aX + Ab$$

なる変換を定義すると

$$G = \{g_{ab}, \quad a \in R^+, \quad b \in R^s\}$$

は $\qquad g_{a,b} \cdot g_{a',b'} = g_{aa', ab'+b}$

なる乗法に従う $R^n$ 上の変換群である。又

$$\mathring{G} = \{\mathring{g}_{a,b}\}, \quad \mathring{g}_{ab} : (\mu, \sigma^2) \longrightarrow (a\mu+b, \ a^2\sigma^2)$$

$\bar{G} = \{\bar{g}_{a,b}\}, \quad \bar{g}_{ab} : \quad \psi \rightarrow a\psi + b$ 　　但し $\mu$ の推定問題を考えるときで $\psi \in R^s$

$\qquad\qquad\qquad\qquad \phi \longrightarrow a^2\phi$ 　　但し $\sigma^2$ の推定を考えるときで $\phi \in R^+$

とすると

$$\hat{\mu}(g_{ab}X) = a\hat{\mu}(x) + b = \bar{g}_{ab}\hat{\mu} \quad , \quad S(g_{ab}X) = a^2 S(x)$$

が成り立つ。この式より最小２乗推定量 $\hat{\mu}(x)$ は $G$－不変である。

$\mu$の推定値を$\psi$とするとき損失関数$L$を

(4) $$L[(\mu,\sigma^2),\psi]=\frac{1}{\sigma^2}(\psi-\mu)'\Sigma^{-1}(\psi-\mu)$$

とおくと，$L$は$G$−不変である。

最小２乗解$\hat{\mu}$のリスクは(2)を用いて

$$\rho[(\mu,\sigma^2),\hat{\mu}]=\frac{1}{\sigma^2}E\left(\hat{\mu}(x)-\mu\right)'\Sigma^{-1}(\hat{\mu}(x)-\mu)=\frac{1}{\sigma^2}E\,\varepsilon' e\,\varepsilon$$
$$=s \qquad \textit{for all}\ (\mu,\sigma^2)\in R^s\oplus R^+$$

$\hat{\mu}(x)$はミニマックス$G$−不変推定量である．

$(\hat{\mu},S)$は最小十分統計量であるから，以下では$\mu$の推定量として$\hat{\mu}$と$S$の関数であるもののみを考える。(5)及び(5')で与えられる$\mu$の推定量を作り，それが$s>2$のときには$\hat{\mu}$の改良になっていることを示す。

(5) $$\varphi(\hat{\mu},S)=\left(1-\frac{aS}{\hat{\mu}'\Sigma^{-1}\hat{\mu}}\right)\hat{\mu}$$

とおく。$\varphi$を$\mu$の推定値としたときのリスクを最小ならしめる様な$a$を定め、且つそのときのリスクを計算しよう．

$$\rho=\rho[(\mu,\sigma^2),\varphi]=\frac{1}{\sigma^2}E\left[\varphi(\hat{\mu},S)-\mu\right]'\Sigma^{-1}\left[\varphi(\hat{\mu},S)-\mu\right]$$

$$=\frac{1}{\sigma^2}E\left[(\hat{\mu}-\mu)'\Sigma^{-1}(\hat{\mu}-\mu)-2aS\frac{\hat{\mu}'\Sigma^{-1}(\hat{\mu}-\mu)}{\hat{\mu}'\Sigma^{-1}\hat{\mu}}+a^2S^2\frac{1}{\hat{\mu}'\Sigma^{-1}\hat{\mu}}\right]$$

$$=s-2a(n-s)E\frac{\hat{\mu}'\Sigma^{-1}(\hat{\mu}-\mu)}{\hat{\mu}'\Sigma^{-1}\hat{\mu}}+a^2(n-s)(n-s+2)E\frac{\sigma^2}{\hat{\mu}'\Sigma^{-1}\hat{\mu}}$$

ここで $\Sigma=BB'$ 但し$B$は$s\times s$三角形行列 $b_{ij}=0,\ i<j$ で対角線は正のもの， なる$B$を用いて．

$$Y = B^{-1}\hat{\mu}$$

なる変換を行うと

$$Y : N(\bar{\mu}, I_s\sigma^2) \qquad 但し \quad \bar{\mu} = B^{-1}\mu$$

このYを用いてρを計算すると

$$\rho = s - 2a(n-s)\,E\,\frac{Y'(Y-\bar{\mu})}{Y'Y} + a^2(n-s)(n-s+2)\,E\,\frac{\sigma^2}{Y'Y}$$

ここで Stein [2] p364 と同様の計算により，Kを平均 $\frac{1}{2\sigma^2}\bar{\mu}'\bar{\mu}$

$= \frac{1}{2\sigma^2}\mu'\Sigma^{-1}\mu$ なるポアッソン変数として

$$\rho = s - 2a(n-s)\,E\,\frac{s-2}{s-2+2K} + a^2(n-s)(n-s+2)\,E\,\frac{1}{s-2+2K}$$

このリスクは

$$a = \frac{s-2}{n-s+2}$$

のとき最小となり，その値は

$$\rho = s - \frac{(n-s)(s-2)^2}{n-s+2}\,E\,\frac{1}{s-2+2K} < s = \rho\left[(\mu, \sigma^2), \hat{\mu}\right]$$

よって最小2乗解は許容的でない。　($s > 2$ のとき).

上と同様の議論により

(5')　$$\varphi_k(\hat{\mu}, s) = \left(1 - \frac{aS}{(\hat{\mu}-k)'\Sigma^{-1}(\hat{\mu}-k)}\right)(\hat{\mu}-k) + k$$

但し $k \in R^s$,　$a = \frac{s-2}{n-s+2}$

なる $\varphi_k$ はすべて $\hat{\mu}$ の改良になっている。

$\hat{\mu}$ も $\varphi_k$ もミニマックス推定量である.

## § 3 Eliptically symmetric estimator

$$\Pi = \{ g : \ g = BQB^{-1}, \ \ \Sigma = BB', \ \ Q \in O_s \ (s\times s \text{ 直交行列}) \}$$

$\mu$ の推定量 $\psi(\hat{\mu}, S)$ が $g \circ \psi \circ g^{-1} = \psi$ をみたすとき *Eliptically symmetric estimator* という．但し $g \circ \psi \circ g^{-1} = \psi$ は

$\forall\, b \in R^s \quad g[\psi(g^{-1}b, S)] = \psi(b, S)$ の意味．

§ 2 の $\varphi(\hat{\mu}, S)$ がこの性質を持っていることは、次の様にしてわかる．

$$g[\varphi(g^{-1}\hat{\mu}, S)] = BQB^{-1}\left(1 - \frac{aS}{\hat{\mu}' B'^{-1} Q' B' \Sigma^{-1} BQB^{-1} \hat{\mu}}\right) BQ'B^{-1}\hat{\mu}$$

$$= \left(1 - \frac{aS}{\hat{\mu}' \Sigma^{-1} \hat{\mu}}\right)\hat{\mu} = \varphi(\hat{\mu}, S)$$

一般に

$$\psi(\hat{\mu}, B) = H(\hat{\mu}'\Sigma^{-1}\hat{\mu}, S)\,\hat{\mu} \qquad \text{但し } H \text{ は実数値関数}$$

する形の $\psi$ は *Eliptically symmetric* である

逆にすべての $g \in \Pi$ につき $g \circ \psi \circ g^{-1} = \psi$ をみたす $\psi$ はどの様な関数であるかということを考えて見る．

$\Pi = \{g\}$ は $\hat{\mu}'\Sigma^{-1}\hat{\mu} = \text{constant}$ 上の点をその上の点に写す変換全体である．原点と $\hat{\mu}$ を結ぶ直線を $L(\hat{\mu})$ とするとき $\Pi$ の中に $L(\hat{\mu})$ のみを不動にする $g$ がある。その $g$ については

$g[\psi(g^{-1}\hat{\mu}, S)] = \psi(\hat{\mu}, S)$ は $g[\psi(\hat{\mu}, S)] = \psi(\hat{\mu}, S)$ となる．故に $\psi(\hat{\mu}, S)$ は $L(\hat{\mu})$ 上の点である．

$\therefore \ \exists\, H = H(\hat{\mu}, S)$ *real valued function*, $\psi(\hat{\mu}, S) = H(\hat{\mu}, S)\hat{\mu}$

$$\psi(\hat{\mu}, S) = g\, H(g^{-1}\hat{\mu}, S)\, g^{-1}\hat{\mu} = H(g^{-1}\hat{\mu}, S)\,\hat{\mu} = H(\hat{\mu}, S)\,\hat{\mu}$$

$$\therefore \quad H(g^{-1}\hat{\mu}, S) = H(\hat{\mu}, S) \qquad \forall g \in \Pi$$

故に H は $\hat{\mu}'\Sigma^{-1}\hat{\mu}$ と S の関数である。則ち

(6) $$\psi(\hat{\mu}, S) = H(\hat{\mu}'\Sigma^{-1}\hat{\mu}, S)\,\hat{\mu}$$

このとき次の事が成立する

$\Psi = \{\psi : \psi = H(\hat{\mu}'\Sigma^{-1}\hat{\mu}, S)\,\hat{\mu}\}$ の中で許容的な推定量は推定量全体 $\Delta$ の中で許容的である。　証明。$\psi$ を $\Psi$ の中では許容的であるが $\Delta$ の中で許容的でないとし $\phi \in \Delta$ を $\psi$ の改良であるとする。則ち

$$\rho[(\mu, \sigma^2), \phi] \leqq \rho[(\mu, \sigma^2), \psi]$$

$$\boxed{\forall (\mu, \sigma^2) \in R^d \oplus R^+}$$

$$\exists\ (\mu_0, \sigma_0^2) \qquad \rho[(\mu_0, \sigma_0^2), \phi] < \rho[(\mu_0, \sigma_0^2), \psi]$$

$\rho$ は $(\mu, \sigma^2)$ につき連続であるから $\Omega$ なる open set 上で下の行の不等式が成立する。

$O_d$ はコンパクトであるから $O_d$ 上の invariant measure の total measure を 1 に出来る　その measure より導かれる $\Pi$ 上の measure を $\lambda$ とする。

$$\phi_1(\hat{\mu}, S) = \int_{\Pi} g\,\phi(g^{-1}\hat{\mu}, S)\, d\lambda(g)$$

により $\phi_1$ を定義すると $\phi_1 \in \Psi$ である

$$\because \quad g_1 \circ \phi_1 \circ g_1^{-1} = g_1 \int_{\Pi} g\,\phi(g^{-1}g_1^{-1}\hat{\mu}, S)\, d\lambda(g) = \phi_1$$

又 $\forall \phi \in \Delta$ につき

(7) $$\rho[(\mu, \sigma^2), g\circ\phi\circ g^{-1}] = \rho[(g^{-1}\mu, \sigma^2), \phi]$$

が成立する

$$\because\ \rho[(\mu,\sigma^2),\ g\circ\phi\circ g^{-1}] = E\{g\phi(g^{-1}\hat{\mu}, s)-\mu\}'\Sigma^{-1}\{g\phi(g^{-1}\hat{\mu}, s)-\mu\}/\sigma^2$$

$$= E\{\phi(g^{-1}\hat{\mu}, s)-g^{-1}\mu\}' g'\Sigma^{-1} g\{\phi(g^{-1}\hat{\mu}, s)-g^{-1}\mu\}/\sigma^2$$

ここで $g'\Sigma^{-1}g = B^{-1\prime}Q'B'B'^{-1}B^{-1}BQB^{-1} = \Sigma^{-1}$ であるので

$$= \rho[(g^{-1}\mu,\ \sigma^2),\ \phi]$$

(8) $(g^{-1}\mu,\ \sigma^2)\in\Omega \rightarrow \rho[(\mu,\sigma^2),\ g\circ\phi\circ g^{-1}] < \rho[(g^{-1}\mu,\sigma^2),\ \psi]$

凸関数 $W$ については $W E(x) \leq E W(x)$ であるから

$$(\phi_1(\hat{\mu}, s)-\mu)'\Sigma^{-1}(\phi_1(\hat{\mu}, s)-\mu)$$

$$\leq \int_{\Pi}\{g\phi(g^{-1}\hat{\mu}, s)-\mu\}'\Sigma^{-1}\{g\phi(g^{-1}\hat{\mu}, s)-\mu\}\,d\lambda(g)$$

両辺に $\hat{\mu}$, $s$ の同時分布の密度関数を掛けて平均すると （積分順序の変更可能）

$$\rho[(\mu,\sigma^2),\ \phi_1] \leq \int_{\Pi}\rho[(\mu,\sigma^2),\ g\circ\phi\circ g^{-1}]\,d\lambda(g)$$

(7) より

$$= \int_{\Pi}\rho[(g^{-1}\mu,\sigma^2),\ \phi]\,d\lambda(g)$$

(8) より

$$< \int\rho[(g^{-1}\mu,\sigma^2),\ \psi]\,d\lambda(g) = \int\rho[(\mu,\sigma^2),\ g\circ\psi\circ g^{-1}]\,d\lambda(g)$$

$$= \int\rho[(\mu,\sigma^2),\ \psi]\,d\lambda(g) = \rho[(\mu,\sigma^2),\ \psi]$$

所が $\psi$ も $\phi_1$ も共に Eliptically symmetric で $\psi$ は その中で 許容であったから これは 矛盾 である.

(6') $\psi = H((\hat{\mu}-k)'\Sigma^{-1}(\hat{\mu}-k))(\hat{\mu}-k)+k \qquad k\in R^s$

なる推定量全体を $\Phi_k$ とかき Eliptically symmetric about $k$ ということにすると上と同様の結果が導かれる.

## §4. Bayes estimator

母数空間上の事前分布として $\sigma^2 = \mathrm{const}$ 上の分布 $N(0,\ BDB'\sigma^2)$ をとる. 但し

$$D = \begin{bmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_s \end{bmatrix} \qquad d_i > 0 \quad \text{で} \quad B \text{は} \ \Sigma = BB'$$

$$M = (I_s + D^{-1})^{-1} = \begin{bmatrix} \frac{1}{1+d_1^{-1}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{1+d_s^{-1}} \end{bmatrix} = \begin{bmatrix} m_1 & & \\ & \ddots & \\ & & m_s \end{bmatrix}, \quad 0 < m_i < 1$$

とおき $\mu$ の推定量としての Bayes 解を求めると

(9) $$\varphi_M(\hat{\mu}, s) = BMB^{-1}\hat{\mu}$$

又事前分布として $N(0,\ BQDQ'B'\sigma^2)$ をとると Bayes 解は

(9') $$\varphi_{M,Q}(\hat{\mu}, s) = BQMQ'B^{-1}\hat{\mu} \qquad \text{但し } Q \in O_s$$

又事前分布として $N(k,\ BQDQ'B'\sigma^2)$ をとると Bayes 解は

(9'') $$\varphi = BQMQ'B^{-1}(\hat{\mu} - k) + k$$

逆に適当な $0 < m_i < 1$ $i = 1, \cdots s$ をとって (9'') を作ってもそれは適当な事前分布に対する unique Bayes 解であるから許容的である.

(9) に対する ベイズ リスク もポめて見ると

$$\rho = \sum_{i=1}^{s} m_i^2 + \frac{1}{\sigma^2}\mu' B'^{-1}\begin{bmatrix}(1-m_1)^2 & & 0\\ & \ddots & \\ 0 & & (1-m_s)^2\end{bmatrix}B^{-1}\mu$$

事前分布で $d_i = 0$ のとき (singular な分布) も考えると

$0 \leqq m_i < 1$

$d_i \to \infty$ のときを考えると $m_i \to 1$ になるが このときは $m_i = 1$ になる $i$ が 2個以上であれば §2 の結果より許容でないことがわかっていて 結局

$0 \leqq m_i \leqq 1$ $i = 1, \cdots s$ で 但し $m_i = 1$ になる $i$ は 2個以下 であるとき (9″) は 許容的 になる.

## §5 多変量正規分布の平均ベクトルの推定

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} : N(\xi, \Sigma), \quad E(X) = \xi, \quad E(X-\xi)(X-\xi)' = \Sigma$$

$p > 2$

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1p} \\ & \ddots & \vdots \\ s_{p1} & & s_{pp} \end{bmatrix} : W(n, \Sigma) \quad \text{ウィッシャート分布}$$

$s_{ij} = s_{ji}$ $\quad E(S) = n\Sigma$

$X$ と $S$ は独立とし $X$ と $S$ を得て $\xi$ の推定する.

損失関数 $L[(\xi, \Sigma), \hat{\xi}] = (\hat{\xi} - \xi)' \Sigma^{-1} (\hat{\xi} - \xi)$

$\xi$ の推定量として $X$ そのものをとったときの リスク は

$$\rho[(\xi, \Sigma), X] = p$$

$\xi$の推定量として

(10) $$\varphi(X, S) = \left(1 - \frac{c}{X'S^{-1}X}\right)X$$

をとったときのリスクは

(11) $C = \dfrac{p-2}{n-p+3}$ のとき最小になりそのとき

$$\rho[(\xi, \Sigma), \varphi] = p - \frac{n-p+1}{n-p+3}(p-2)^2 E\frac{1}{p-2+2K}$$

K はポアッソン変数で平均 $\frac{1}{2}\xi'\Sigma^{-1}\xi$ であるもの.

(10) に対するリスクを計算する為に

$\Sigma = BB'$ : B は三角形行列 $b_{ij}=0$ , $j>i$ . $b_{ii}>0$

$Y = B^{-1}X$ : $N(\bar{\xi}, I_p)$ $\bar{\xi} = B^{-1}\xi$

$\bar{S} = B^{-1}SB'^{-1}$ : $W(n, I_p)$

とおいて

$$\rho[(\xi, \Sigma), \varphi] = E_{\xi,\Sigma}\{\varphi(X,S)-\xi\}'\Sigma^{-1}\{\varphi(X,S)-\xi\}$$

$$= E_{\xi,\Sigma}\left\{\left(1-\frac{c}{X'S^{-1}X}\right)X-\xi\right\}'\Sigma^{-1}\left\{\left(1-\frac{c}{X'S^{-1}X}\right)X-\xi\right\}$$

$$= E_{\bar{\xi},I_p}\left\{\left(1-\frac{c}{Y'\bar{S}^{-1}Y}\right)Y-\bar{\xi}\right\}'\left\{\left(1-\frac{c}{Y'\bar{S}^{-1}Y}\right)Y-\bar{\xi}\right\}$$

$Q \in O_p$ を第1行が $\dfrac{1}{\sqrt{\bar{\xi}'\bar{\xi}}}\bar{\xi}'$ であるものとする

$Z = QY$ とおくと $E(Z) = [\sqrt{\bar{\xi}'\bar{\xi}}, 0 \cdots 0]' = \xi^*$ とおく

$\bar{\bar{S}} = Q'\bar{S}Q$ とおく

$Z$ : $N(\xi^*, I_p)$ , $\bar{\bar{S}}$ : $W(n, I_p)$

$$\rho = E_{\xi^*\, I_p}\left\{\left(1-\frac{c}{Z'\bar{\bar{S}}^{-1}Z}\right)Z-\xi^*\right\}'\left\{\left(1-\frac{c}{Z'\bar{\bar{S}}^{-1}Z}\right)Z-\xi^*\right\}$$

$\dfrac{1}{\sqrt{Z'Z}}Z'$ を第1行にもつ $O_p$ の element を $R$ とする

$\bar{V} = RZ$ とおく　$\bar{V} = [\sqrt{Z'Z}, 0, \cdots 0]'$

$\bar{\bar{S}} = R\bar{S}R'$ : $\bar{W}(n, I_p)$　$\bar{\bar{S}}$ の Bartlett decomposition

を $\bar{\bar{S}} = CC'$ とする。

$$\bar{\bar{S}} = \begin{bmatrix} c_{11} & \cdots & c_{1p} \\ & \ddots & \\ 0 & & c_{pp} \end{bmatrix}\begin{bmatrix} c_{11} & & 0 \\ & \ddots & \\ c_{1p} & & c_{pp} \end{bmatrix}$$

$c_{11}$ : d.f. $= n-p+1$ のカイ分布
⋮
$c_{pp}$ : d.f. $= n$ のカイ分布
$c_{ij}$ : $N(0,1)$

$$Z'\bar{\bar{S}}^{-1}Z = \bar{V}'R'\bar{\bar{S}}^{-1}R\bar{V} = \bar{V}'\bar{\bar{S}}^{-1}\bar{V}$$

$$= [\sqrt{Z'Z}, 0 \cdots 0]\begin{bmatrix} c_{11}^{-1} & & 0 \\ \vdots & \ddots & \\ \cdots & \cdots & \end{bmatrix}\begin{bmatrix} c_{11}^{-1} & \cdots & \vdots \\ & \ddots & \vdots \\ 0 & & \end{bmatrix}\begin{bmatrix} \sqrt{Z'Z} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \frac{Z'Z}{c_{11}^2} = \frac{Z'Z}{S^*}$$

$S^* = c_{11}^2$ は d.f. $= n-p+1$ のカイ2乗分布で且 $Z$ と indep.

$$\rho = E^{Z}_{\xi^*, I_p}\ E^{\chi^2}_{n-p+1}\left\{\left(1-\frac{cS^*}{Z'Z}\right)Z-\xi^*\right\}'\left\{\left(1-\frac{cS^*}{Z'Z}\right)Z-\xi^*\right\}$$

これは §2 の(5)で $S = S^*$　$\hat{\mu} = Z$, $\Sigma = I_p$ としてのリスクを計算することになり (11)の結果をうる。

(11)より $p > 2$ のとき $\xi$ の推定量として $X$ は許容的ではない

( $X$ は $\underbrace{\text{不変推定量}}_{\text{ミニマックス}}$ )

## §6. 多変量正規分布の共分散行列の推定

$X_i$ $i=1, 2, \cdots$ $n>p$ を $N(0, \Sigma)$ よりの$n$個の独立標本とし、それより $\Sigma$ を推定する。($X_i$ は $p$次元 $\Sigma$ は $p\times p$ とする)

(12) $$S=\sum_{i=1}^{n} X_i X_i' \quad : \quad W(n, \Sigma)$$

は十分統計量であるから $\Sigma$ の推定量としては $S$ の関数のみを考える。$\Sigma$ の推定量を $\hat{\Sigma}$ とするときどの様な損失関数が適当であるかということについては議論もあるかと思われるが、ここでは

(13) $$L(\Sigma, \hat{\Sigma}) = \operatorname{tr} \Sigma^{-1}\hat{\Sigma} - \log \det \Sigma^{-1}\hat{\Sigma} - p$$

を採用する。

$$G=\{g : g \text{ は } p\times p \text{ nonsingular 行列}\}$$

とし

$$g : \quad X \to gX$$
$$\qquad S \to gSg'$$

なる変換を考えると問題は $G$-不変であり、$\Sigma$ の推定量 $\varphi(S)$ が $G$-不変であるとは

(14) $$\varphi(gSg') = g\,\varphi(S)\,g'$$

をみたすことであり $S$ の常数倍であるところの普通の推定量は不変な推定量である。以下で普通の推定量はミニマックスではないことを示す。

$$A=\{a : a \text{ は } p\times p \text{ 三角行列 } a_{ij}=0 \text{ for } j>i\}$$

$A$-invariant なものの中でミニマックスなものをもとめて見る

(14) で $S=I$ とおき

$$\varphi(aa') = a\,\varphi(I)\,a' \qquad a \in A$$

ここで特に $a$ として、対角形で対角要素が $\pm 1$ のものをとると

$$\varphi(I) = a\,\varphi(I)\,a'$$

$\varphi(I)$ は $p\times p$ 行列であるが 1ヶ所でも対角要素以外の所に 0 でない要素があればこの式は成り立たないから

$$\varphi(I) = D = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_p \end{bmatrix}$$

$$S = KK' \qquad (K \text{ は三角行列で対角線が正})$$

と分けて

(15) $$\varphi(S) = \varphi(KK') = K\,\varphi(I)\,K' = KDK'$$

この $\varphi$ に対するリスクを計算しそれが最小になる様に $D$ を求める。

$$\rho(\Sigma, \varphi) = E[\operatorname{tr} \Sigma^{-1}\varphi(S) - \log\det \Sigma^{-1}\varphi(S) - p]$$

$$= E[\operatorname{tr} \varphi(B^{-1}SB'^{-1}) - \log\det \varphi(B^{-1}SB'^{-1}) - p]$$

但し $\Sigma = BB'$　$B$ は三角行列

$B^{-1}SB'^{-1}$ : $W(n, I_p)$

$B^{-1}SB'^{-1}$ を改めて $S$ とおくと

$$\rho = \rho(I, \varphi) = E[\operatorname{tr} \varphi(S) - \log\det \varphi(S) - p]$$

$$= E[\operatorname{tr} KDK' - \log\det K\Delta K' - p]$$

$$= E\operatorname{tr} KDK' - \log\det \Delta - E\log\det S - p$$

$$E\ \mathrm{tr}\ KDK' = \sum_{i,k} d_i E K_{ki}^2 = \sum d_i E \chi^2_{n-i+1+p-i}$$

$$= \sum d_i (n+p-2i+1)$$

∵ $S : W(n, I_p)$ で $S = KK'$ であるから

$K$ は $S$ の Bartlett の decomposition になっている

$$E \log \det S = \sum_{i=1}^{p} E \log \chi^2_{n-i+1}$$

$$\rho = \sum_{i=1}^{p} [(n+p-2i+1)d_i - \log d_i] - \sum_{i=1}^{p} E \log \chi^2_{n-i+1} - p$$

これは

(16) $$d_i = \frac{1}{n+p-2i+1}$$

にとったとき最小になり その値は

(17) $$\rho(\Sigma, \varphi^*) = \sum_{i=1}^{p}\left[1 - \log\frac{1}{n+p-2i+1} - E \log \chi^2_{n-i+1}\right] - p$$

$$= \sum_{i=1}^{p}\left[\log(n+p-2i+1) - E \log \chi^2_{n-i+1}\right] \qquad (n>p)$$

となる.

(15),(16) で与えられる $\varphi(S)$ がミニマックス推定量である

それが $A$ なる変換群に関係していることは明かで $(1, 2 \cdots p)$ の

適当な permutation を作りその order に従って作った三角行列

全体を $A_\alpha$ とせば リスクの値 (17) は同じであるが推定量

としては異なる $\varphi(S)$ をうる

§2,3 で $A = I_n$ とおくと Stein [1] の §2, §3 及び Stein [2] の §2の一部

§5.6 は Stein [2] の §2, §5 と全く同じ

§4 で $A = I_n$ とおくと Cohen Ann. Math Stat. 1966